Computational Accelerator Physics Grand Challenge: Summary of FY99 Activities

Robert D. Ryne[1] and Kwok Ko[2]

with contributions from

James Ahrens[5], Timothy Cleland[5], Julian Cummings[5], Viktor Decyk[4], Nathan Folwell[2], Gene Golub[3], Salman Habib[1], William Humphrey[5], Steven Karmesin[5], Patrick McCormick[5], Allen McPherson[5], Ji Qiang[1], Zenghai Li[2], Brian McCandless[2], Graham Mark[5], Wanjun Mi[3], Cho-Kuen Ng[2], Mikhail Saparov[3], William Saphir[6], Yong Sun[3], Michael Wolf[2]

[1]Los Alamos National Laboratory
[2]Stanford Linear Accelerator Center
[3]Stanford University
[4]University of California, Los Angeles
[5]Advanced Computing Laboratory
[6]National Energy Research Scientific Computing Center

September 24, 1999

1 Introduction

The goal of the Grand Challenge in Computational Accelerator Physics is to develop a new generation of accelerator simulation tools, targeted to high performance computing platforms, and to apply them to problems of national importance. In so doing, we aim to have a major impact on the design, performance optimization, and ultimately, on the success of next-generation accelerators. Such accelerators will push the frontiers of beam energy, beam intensity, beam quality, and accelerator system complexity. The design of major accelerator facilities involves an enormous investment in theory, experiment, and simulation. Neglecting any of these can lead to cost overruns, degraded performance, inability to meet performance requirements, and ultimately, the failure of major projects. The development of a terascale accelerator modeling capability will lead to increased confidence in designs and cost optimizations that may save 100's of millions of dollars in construction and operating costs. The Grand Challenge has already resulted in a 1000-fold increase in accelerator simulation capability, and tools developed under the Grand Challenge are now being used to model existing and planned accelerator facilities.

The activities of the Grand Challenge fall into two main categories: Beam Dynamics and Electromagnetics. The Beam Dynamics effort is aimed at developing a high-resolution particle simulation capability that is needed to design the next-generation of high intensity accelerators. The goal is to develop a capability that will enable *quantitative* predictions of the beam's evolution, including the tails of the distribution (the so-called "beam halo"), as the beam propagates through kilometers of accelerator structures and components. A major accomplishment of this effort is IMPACT, a code that combines modern methods of particle-in-cell simulation with nonlinear Magnetic Optics, all in a High Performance Computing (HPC) environment. IMPACT has already been used in support of the Spallation Neutron Source (SNS) and Accelerator Production of Tritium projects. The Electromagnetics effort is aimed at the development of a suite of parallel field solvers capable of modeling complex beamline components and systems using terascale computing resources. This suite will have a comprehensive set of capabilities that will enable an accelerator physicist/engineer to solve, with unprecedented accuracy, a wide range of electromagnetic problems that arise in accelerator design and analysis. Two major accomplishments under the Grand Challenge are the eigensolver Omega3P and the time-domain solver Tau3P. Omega3P solves for the normal modes of RF cavities while Tau3P calculates the transmission properties of structures. These codes have already been used to model electromagnetic structures for the SNS project, for the Next Linear Collider (NLC), and for the Advanced Light Source (ALS).

Below we summarize our activities for FY99 and our plans for FY2000.

2 Beam Dynamics

2.1 Development of IMPACT

The availability of parallel supercomputers is making it possible to perform beam dynamics simulations with unprecedented speed and resolution. Under the Grand Challenge, we have reprogrammed existing codes and have developed new codes to take advantage of parallel computing resources. Our emphasis to-date has been in modeling high current linear accelerators, and our primary code for modeling linacs on parallel computers is IMPACT (Integrated-Map and Particle Accelerator Tracking code). IMPACT combines techniques of Magnetic Optics with those of parallel Particle-In-Cell (PIC) simulation by using split-operator methods. In this approach, the mean-field Hamiltonian governing the dynamics of particles is separated into two pieces, $H=H_{ext}+H_{sc}$, where H_{ext} corresponds to externally applied fields and H_{sc} corresponds to space-charge fields. The effect of H_{ext} is treated using map-based techniques of Magnetic Optics, while the effect of H_{sc} is treated by solving Poisson's equation on a 3D grid using an FFT-based convolution algorithm. IMPACT has the following features:

- Parallel implementation using F90/MPI, POOMA, and HPF. The F90/MPI and POOMA versions both use "particle managers" and have dynamic load balancing.
- Space charge calculation using a 3D (x-y-z) Poisson solver with open boundary conditions. The code assumes the fields are electrostatic in the bunch frame. Relativistic effects (i.e. reduction in the longitudinal force due to the azimuthal magnetic field of the bunch) are included.
- Use of canonical variables, (x,p_x,y,p_y,t,p_t), where t is time-of-flight with respect to the reference particle (equivalent to a phase when normalized), and p_t is the (negative) energy deviation with respect to the reference particle.
- Numerical integration of the reference particle trajectory during the simulation.
- Beamline elements (drift spaces, magnetic quadrupoles, and rf accelerating gaps) are treated using transfer map techniques.
- The transfer matrix for rf accelerating gaps is found by integrating the equations of motion for the map itself, using rf gap fields obtained from the code SUPERFISH. This approach provides for greater accuracy than methods used in other codes.

This year the POOMA version of IMPACT was used during a sustained 48 hour run on 2048 processors of the Nirvana system at LANL as part of the Nirvana acceptance test. This was a tour-de-force in system stability, sheer computing power, and realism in simulation. We anticipate that IMPACT will begin running on the NERSC-3 IBM SP during the latter part of its availability test. IMPACT has also been selected as part of the suite of test codes to be used for the procurement of a future 30 T-Op system at LANL.

Thanks to the Grand Challenge, for the first time we are approaching real-world charge resolution in our beam dynamics simulations. For example, we have run 500 million particle simulations of the Spallation Neutron Source (SNS) linac on the Nirvana system, and the real number of particles per bunch in the SNS linac will be 900 million. IMPACT has also been used to model the Accelerator Production of Tritium linac (which would have 2 billion particles/bunch).

As an example of our capability, we now present results of IMPACT simulations of the SNS Coupled Cavity Linac (i.e. the portion of the linac from 79 MeV to 1000 MeV). Prior to the use of HPC resources, all SNS beam dynamics simulations were two-dimensional and performed on PCs. Typical runs used 10,000 simulation particles, although runs using up to 1 million particles (requiring roughly 2 days to complete) were also performed. It was found that, in runs going from 10,000 to 100,000 to 1,000,000 simulation particles, the maximum amplitude of particles in the simulations steadily increased. Predicting the maximum amplitude is of prime importance, because it provides guidance in choosing the beam pipe aperture. This has enormous consequences: a smaller aperture leads to reduced costs, but it also leads to increased risk of beam interception and resulting radioactivation. Figure 1 shows the maximum particle amplitude as a function of beam energy based on IMPACT runs using 1 million, 10 million, 100 million, 200 million, and 500 million particles. For the 500 million particle case, the charge density was stored on a grid of size 256^3 (corresponding to a 512^3 computational grid that was used, along with a modified Green function, to rigorously treat open boundary conditions). It is seen that the increase in maximum amplitude is approaching a limiting value with runs above 100 million particles. It is worth pointing out that a run using 100 million particles requires 5 hours on 256 processors of the Nirvana system. Compared with a PC run of 1 million particles which required 2 days, our 5 hour parallel run, which is 100 times bigger and 10 times faster than the PC run, represents in increase in simulation capability of a factor of 1000. Furthermore, the PC run was two-dimensional, while the IMPACT run was three-dimensional.

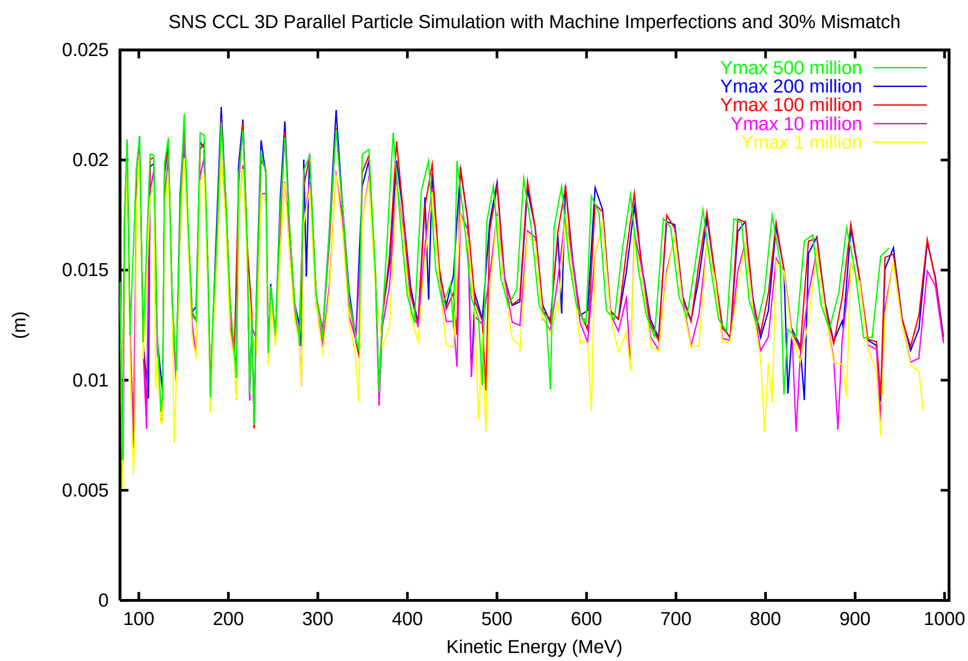

Figure 1: Maximum particle amplitude in the Spallation Neutron Source linac with varying # of particles, simulated on the ACL Nirvana system using 32-1024 processors.

Several improvements were made to IMPACT in FY99. Of particular importance, we achieved a 4x increase in the performance of the F90/MPI version due to a replacement of the original charge deposition/field interpolation routines (based on an algorithm due to Ferrell and Bertschinger) with a parallel particle manager, and through the inclusion of dynamic load balancing. Using the methodology developed by our UCLA collaborators, a 3D object-oriented version of the code was developed, and comparisons were made with other PIC codes used in the plasma community. Other improvements for FY99 include significantly reduced memory overhead, a choice of parallel particle managers (with fixed and variable message buffers), parallel I/O, and checkpoint/restart capabilities. Furthermore, the POOMA version was modified to improve the performance of FFT's across boxes on the SGI Origin 2000 system. In FY99 IMPACT was used to model the APT and SNS linacs. In order to model more realistic accelerators, we added the capability to include machine imperfections. Also, in FY99 IMPACT was used in the first systematic study of halo formation due to longitudinal/transverse coupling in intense charged particle beams (see Figure 2). IMPACT was also used to perform the largest simulations to-date of the SNS linac, with 500 million particles.

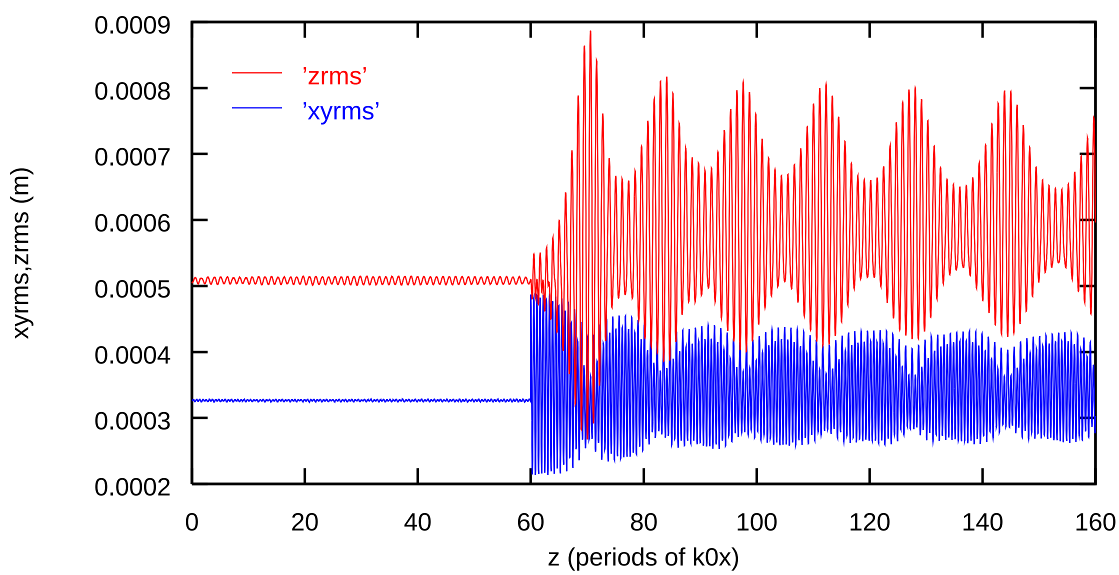

Figure 2: Output from a multiparticle simulation performed on the NERSC T3E showing excitation of a longitudinal mode by an initial transverse excitation applied at z=60. The simulation results validate analytic calculations, which predict strong coupling at the parameters chosen, for which the transverse eigenfrequency is twice that of the longitudinal eigenfrequency (obtained from a linearized eigenmode analysis). [From "Coherent Coupling Criterion for Three-Dimensional Halo Formation," presented at the 1999 International Particle Accelerator Conference.]

Figure 3 shows results of an IMPACT scaling study on the NERSC T3E and the ACL Nirvana Origin 2000 system. For this small problem (2.6 million particles on a 64x64x64 grid), the code scales well. We have also examined different domain decompositions. Figure 4 shows the execution time as a function of the number of processors on the T3E for 1D and 2D partitioning schemes for the same size problem as in Figure 3.

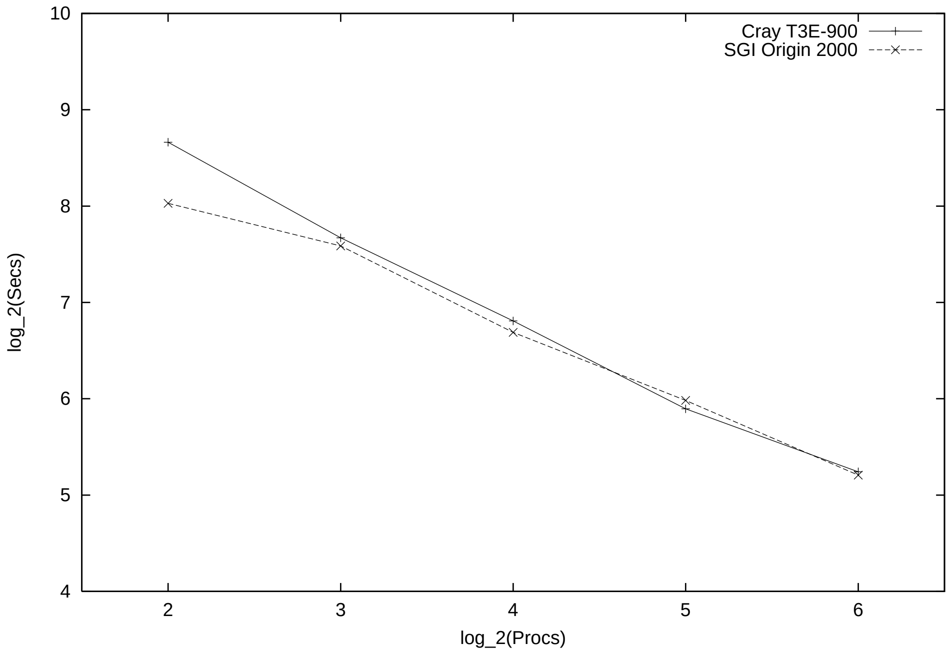

Figure 3: Scaling results on the NERSC T3E and the ACL/Nirvana Origin 2000.

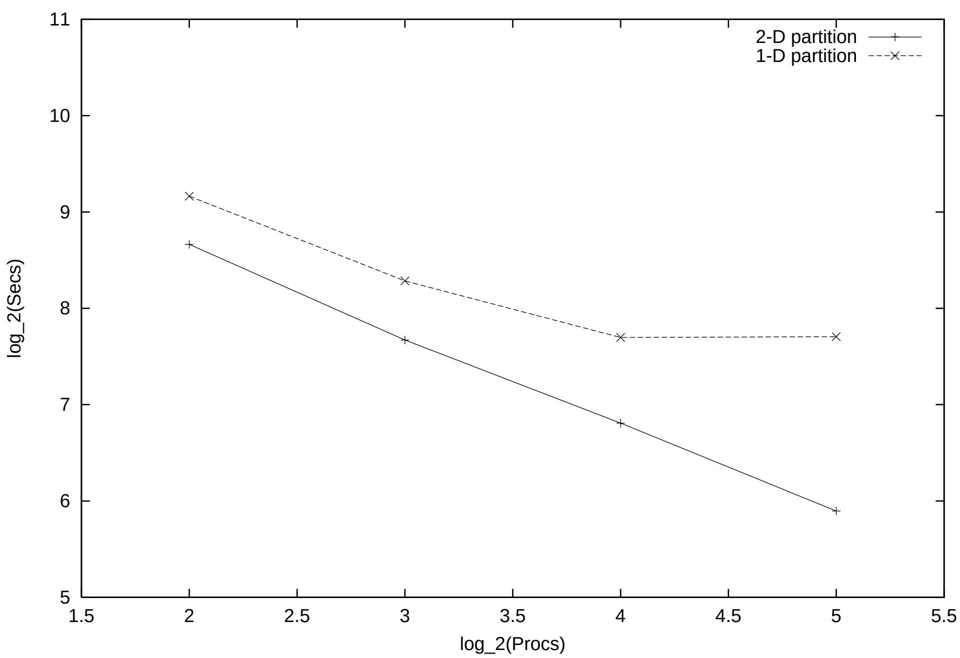

Figure 4: Scaling results on the NERSC T3E for 1D and 2D domain decomposition.

2.2 Very Large Scale Simulations

Thanks to the availability and stability of the 2048 node Nirvana cluster at the ACL, the size of our typical simulations increased dramatically in FY99. We now routinely run 100 million particle simulations across two SMP nodes (i.e. on 256 processors) of the cluster. We have also begun to perform much larger simulations, as exemplified by the 500 million particle simulation (run on 1024 processors) shown in Figure 1. We are now in the process of performing detailed scaling studies to determine the variation in performance across multiple SMP nodes. The ability to run efficiently across nodes is of great importance since it appears that, in the near and foreseeable future, very large scale computing will be performed on SMP clusters. Up to now we have rarely run jobs on more than 128 processors at NERSC, but with the arrival of the much larger IBM/SP we anticipate a similar increase in job size there in FY2000.

2.3 Visualization

In FY99 we began to make greater use of high performance visualization resources. As an example of the amount of data that we must deal with, consider a 500 million particle simulation. The particle array itself is of size 6 x 500,000,000 (since there is a 6-vector (x,p_x,y,p_y,t,p_t) for each simulation particle). Stored as an array of 4 byte quantities, the amount of data generated at each step is 12 Gbyte. A typical simulation contains several thousand steps, corresponding to a total output of several 10's of Tbytes. In order to extract the most useful information, we must also deal with the fact that the data are 6-dimensional. We typically visualize and analyze several 3D projections of our data, for example (x,y,z), (x,p_x,z), and (x,t,p_t). The first of these shows the beam density in real space; the second shows the transverse (horizontal) phase space as a function of longitudinal position within the beam bunch; and the third shows the longitudinal phase space as a function of transverse position within a bunch.

Working closely with the visualization team at the Advanced Computing Laboratory we are now using volume rendering techniques to look at our simulation results. In this approach, a 3D scalar field is treated as a semi-transparent medium. Color and transparency are controlled by scalar transfer functions, i.e. by "color" and "alpha" maps, respectively. The Infinite Reality (IR) graphics pipes on the Nirvana system at the ACL each have 64 Mbyte of texture memory and can hold up to a 512x256x256 volume. A single pipe is able to render volumes of size 256^3 at interactive rates. However, we are now rendering 512^3 data sets. This is accomplished by using multiple IR pipes, and subdividing the volume into sub-volumes, with each sub-volume handled by a separate pipe. Figure 5 shows images based on simulations performed for the Spallation Neutron Source project. The data being visualized is the 3D projection (x,p_x,z). The image on the left his more transparency in order to make it possible to examine the interior structure. The image on the right has less transparency in order to make more visible the largest-amplitude particles, which have the greatest likelihood of intercepting the beam pipe.

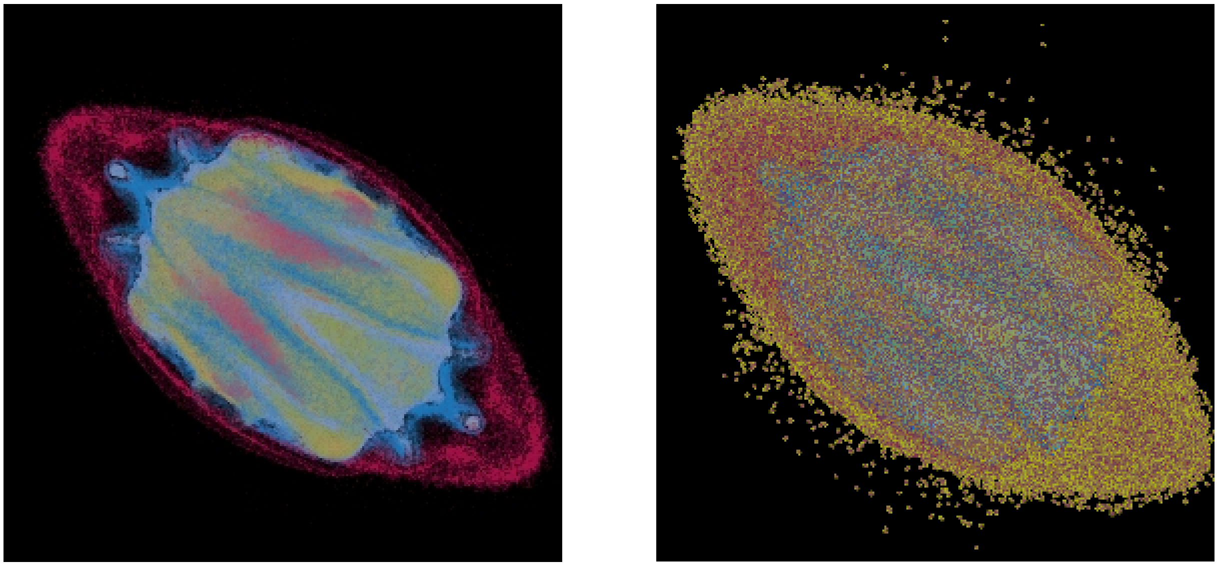

Figure 5: Volume rendering of output from a simulation of the Spallation Neutron Source

Figure 6 is based on the same simulations as Figure 5, but the data are being seen from a different viewpoint in order to exhibit other parts of the 3D projection. Different color maps and alpha maps are also being used. The figure on the left has a high degree of opacity and has the z-axis coming out of the paper, which provides for a clearer view of the (x-p_x) phase space, including its s-shaped tail produced by amplitude-dependent phase advance. The figure on the right is from a different viewpoint and shows how, using transparency, we can look at what are effectively 4 separate isosurfaces of density.

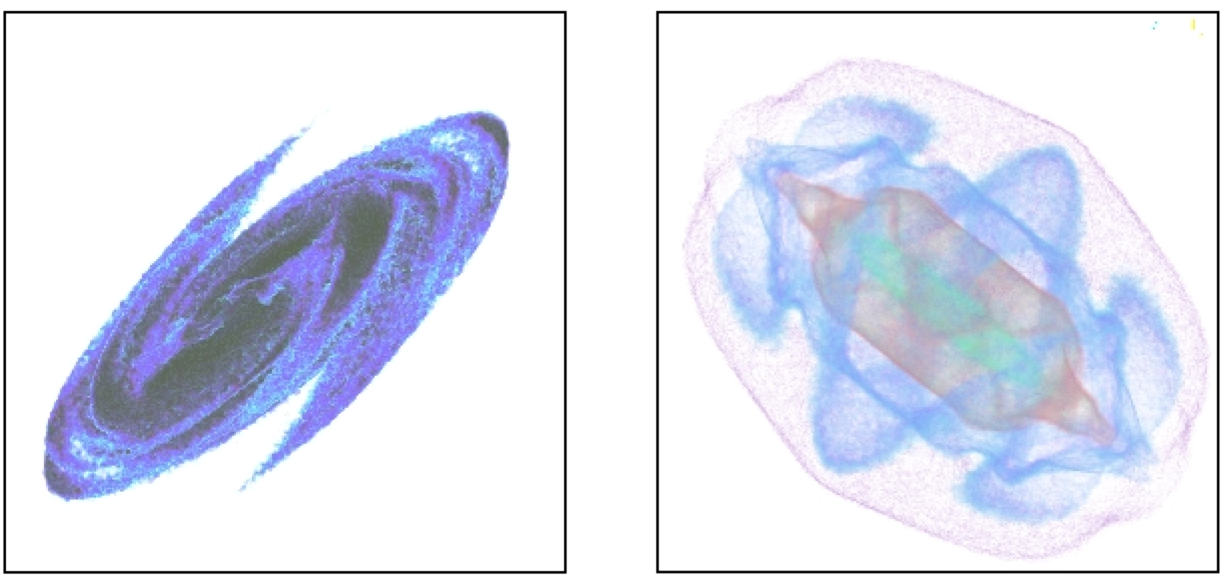

Figure 6: Volume rendering of output from a simulation of the Spallation Neutron Source linac, shown from different viewpoints and with varying degrees of transparency.

2.4 Other Beam Dynamics Activities

Besides IMPACT we have developed two other parallel beam dynamics codes, HALO3D and LANGEVIN3D. Like IMPACT, both codes utilize split-operator methods. HALO3D, developed in collaboration with the University of Maryland, is used for fundamental studies of beam halo phenomena. It includes a capability to initialize certain 3D nonlinear stationary solutions of the Vlasov equation. LANGEVIN3D solves the Fokker-Planck equation with fixed damping and diffusion coefficients. It has been used in simulations to drive a beam artificially to thermal equilibrium prior to mismatching it. The capabilities of HALO3D and LANGEVIN3D are useful when studying halo formation, because nonstationary distributions undergo relaxation after the start of a simulation, often masking other halo-forming mechanisms.

Through a collaborative effort involving SLAC and NERSC, in FY99 we helped initiate and are now using a 34 processor PC cluster located at the SLAC Computing Center. We recently succeeded in porting an ion tracking program to the cluster that is important to the NLC project. We also ported the eigenmode code Omega3P (see below), which runs on the cluster with a performance within 25% that of the T3E for problems running on up to 16 processors.

Other recent Beam Dynamics activities include (1) development of a new, 2nd-order accurate stochastic integration algorithm which will be used in the coming year to model systems with noise; (2) parallelization of the Lie algebraic beam transport code MaryLie using F90/MPI (which we had previously ported to the T3E using HPF); (3) initial work on a time-based (as opposed to z-based) beam dynamics code in C++/MPI; (4) initial work on a parallel code for modeling Radio Frequency Quadrupoles; and (5) continued collaboration with theorists studying beam halo phenomena, and publication of our results in conference proceedings and refereed journals.

3 Electromagnetics

The increasingly demanding design requirements of next-generation particle accelerators such as the Next Linear Collider (NLC) have placed heavy emphasis on the accuracy and reliability of electromagnetics software. Accelerator components for such next-generation machines must be modeled and analyzed with greater speed, accuracy, and confidence than has previously been possible. Furthermore, great demands will be placed on the software in terms of the size and complexity of the problems that can be modeled. Presently popular electromagnetics codes are inefficient in handling complex geometric shapes, or are limited in their ability to solve large-scale problems. One of the main goals of the Grand Challenge is to address these issues by using unstructured grids and multi-processing capability. Below we will describe two major successes of the Grand Challenge: (1) The development of the parallel eigenmode code Omega3P and (2) the development of the parallel time-domain code Tau3P. Both have the following features:

1. C++ implementation
2. Parallelization using MPI
3. Reuse of existing parallel libraries (namely, ParMETIS and AZTEC)

4. Use of unstructured girds
5. New solvers for fast convergence and scalability
6. Adaptive mesh refinement to improve accuracy and performance

3.1 Development of Omega3P

During FY1999, Omega3P was used to model a 5 million degree-of-freedom cavity problem on the T3E, while half-million degree-of-freedom problems were routinely solved on the SLAC Linux Cluster. Quadratic elements were implemented and a Python interface was written to view mesh quality. A new algorithm to find complex eigenvalues was derived and the parallel version is under development. Furthermore, threaded libraries were written for matrix operations to be used for computations on SMP platforms. The comparison between simulation results and measured data has been excellent, so that Omega3P is now viewed to be a reliable tool for designing complex cavities to high accuracy when used with HPC resources like those at NERSC. Cavity modeling with Omega3P was done for the NLC, SNS, ALS, and the APT accelerator projects. The time-domain solver Tau3P was completely parallelized and now shares the same mesh distribution front-end as Omega3P. The port boundary conditions were improved to give better accuracy in transmission properties. Also, we began work on implementing a rigid beam model in the code. Tau3P has been applied to the NLC, W-Band accelerators, and other programs. Descriptions of Omega3P and Tau3P follow.

Omega3P is a parallel distributed-memory finite-element code for electromagnetic modeling of large, complex 3D structures in the frequency domain. The problem is challenging because the distributed mesh operations are communication intensive, and the parallel eigensolver is computationally expensive. Omega3P is actually the culmination of an effort that began with the development of the serial codes Omega2 and Omega3 (2D and 3D codes, respectively). Following the successful development of these serial codes, we developed the software components needed to assemble a robust, maintainable, and extensible capability for eigenmode modeling of electromagnetic structures on parallel computers. The software components consist of:

1. **Omega3P:** The top-level application module layered on DistMesh and EigenSolver (see below), which contains the finite element formulation and post-processing. It uses MeshTV (from LLNL) for visualization.
2. **DistMesh:** A library for operating on distributed unstructured meshes. Operations include parallel file I/O, partitioning, distribution, global numbering, and refinement. DistMesh uses ParMETIS (from U. Minnesota) for partitioning.
3. **EigenSolver:** Code for solving the generalized eigenvalue problem using a hybrid algorithm. It uses AZTEC (from Sandia National Laboratory) for solving sparse linear systems in parallel.

An example of the type of problem to addressed by Omega3P is shown in Figure 7. The figure shows one octant of 1.5 cells of a Damped Detuned Structure (DDS), the accelerating structure being designed for the NLC. This example contains approximately 460,000 elements and 536,000 degrees of freedom. The figure on the left-hand-side shows the computational mesh, and the figure on the right-hand-side shows a partitioning of the problem into 16 roughly-equal pieces. Figure 8 shows the performance of Omega3P for this problem running on up to 128 processors of the T3E at NERSC.

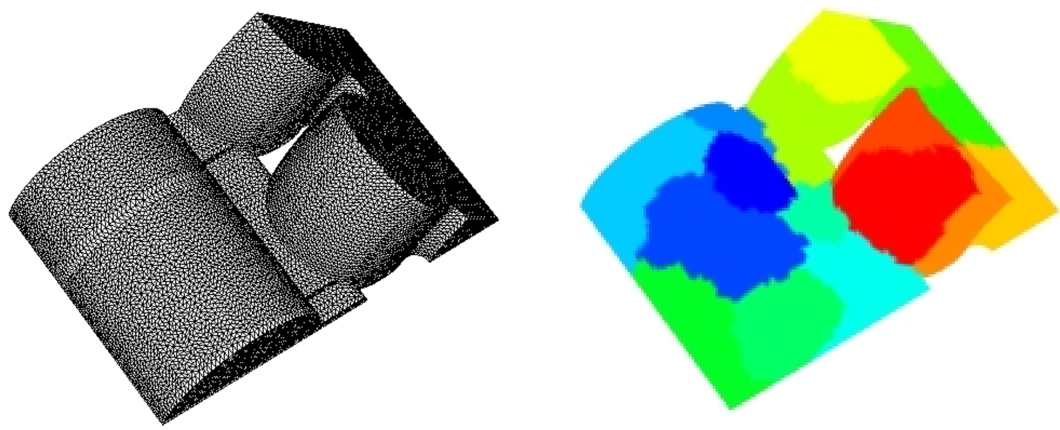

Figure 7: One octant of 1.5 cells of a Damped Detuned Structure (DDS) with a mesh containing 462687 elements, and a partitioning of the problem into 16 pieces.

During the past year we developed, in collaboration with our partners in the Scientific Computing and Computational Mathematics (SCCM) program at Stanford University, a hybrid eigensolver that has excellent performance. It is based on an optimized

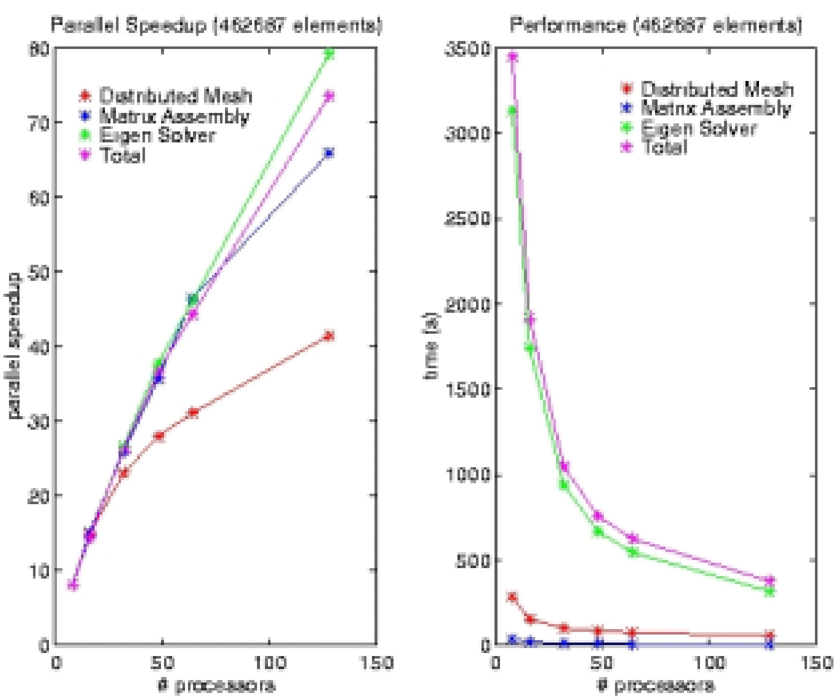

Figure 8: Performance on the NERSC T3E for the problem shown in Figure 7, as a function of the number of processors.

combination involving three powerful methods: (1) spectrum transformation and band pass polynomial filtering, (2) Inexact Krylov subspace projection, and (3) Modified Jacobi-Davidson local refinement. This combination results in solver accuracy and convergence, as well as parallel scalability, far superior to other algorithms that we have utilized. Returning to the problem shown in Figures 7 and 8, Figure 9 shows the eigensolver convergence as a function of the number of steps. For this example, an inexact Krylov subspace projection algorithm was used for six steps, followed by two steps of the Jacobi-Davidson method. It is clear from the Figure that to have continued with the Krylov method would have yielded slow convergence, but by switching to Jacobi-Davidson we greatly accelerate the convergence. Furthermore, we found that the Jacobi-Davidson alone could not lead to rapid convergence without first getting close to the solution by starting with the Krylov method.

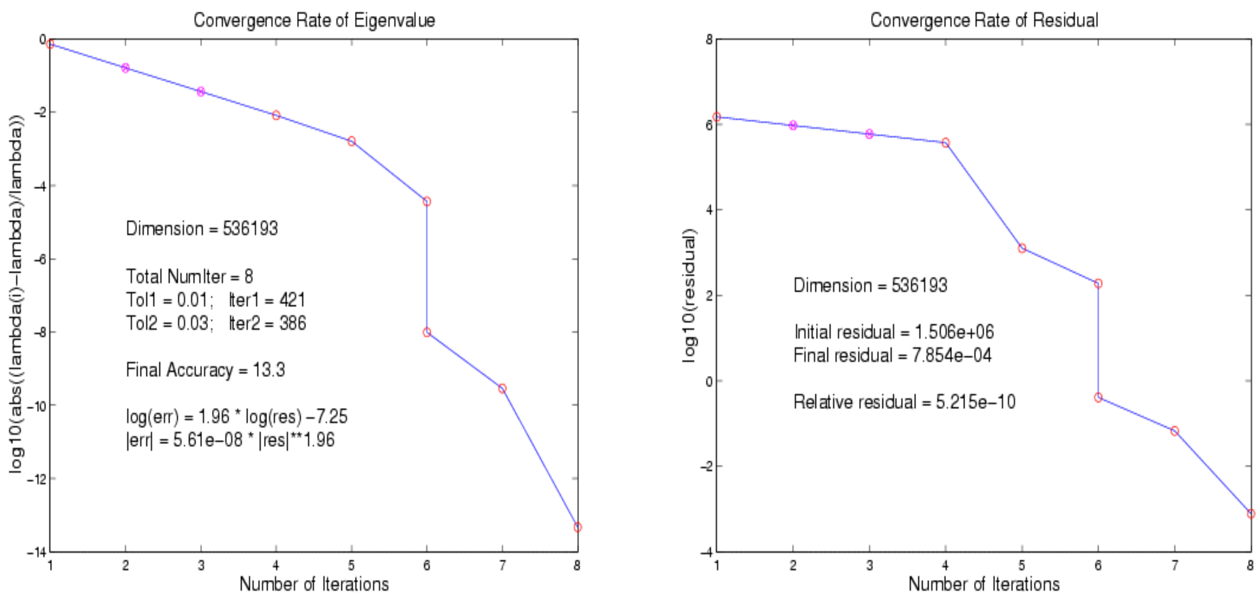

Figure 9: Convergence of the hybrid eigensolver for the problem shown in Figure 7.

3.2 Development of Tau3P

Another major success of the Grand Challenge is the development of Tau3P. This code compliments the capabilities of Omega3P by providing solutions to electromagnetics problems in the time domain (as opposed to the frequency domain). Tau3P is based on a modified Yee algorithm formulated on an unstructured grid. The distribution of the mesh data onto processors is handled by the DistMesh class library as in Omega3P. Tau3P uses a discrete surface integral method to solve Maxwell’s equations, and a leapfrog time-advance scheme with filtering. Since the execution of Tau3P is predominantly matrix-vector operations during time advancement, the program can be readily implemented on distributed memory machines using MPI or on shared memory systems running threads.

Tau3P implements a broadband matched-impedance boundary condition at the waveguide ports to allow for pulse transmission so that S-parameters (such as reflection and transmission coefficients) can be calculated over a frequency range in a *single* simulation. Without this capability, such calculations require a sequence of many runs over a range of frequencies. An example of the type of problem to be addressed by Tau3P is shown in Figure 10. The electromagnetic structure being modeled is an X-band RF choke-cavity used for pulse heating experiments. The long taper for this problem makes it ill-suited to the used of structured grids. As seen from the figure, there is good agreement between measurement and simulation in regard to the bandwidth of the structure. Also, the simulation results are insensitive to changes in the number of elements. The small disagreement between simulation and measurement is likely due to errors in fabrication (i.e. the dimensions of the structure are slightly different from that which was modeled).

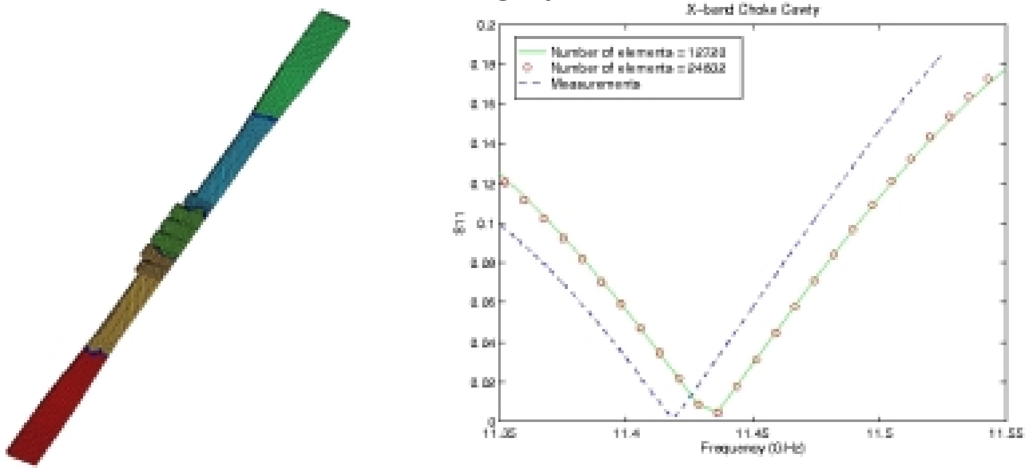

Figure 10: Tau3P results from a simulation of a tapered, X-band RF choke.

Another application of Tau3P is in the design and optimization of couplers for the transmission of rf power into or out of rf structures. An example is shown in Figure 11, in which the reflection coefficient for an X-band input coupler for an NLC accelerator section is found to have a value of 0.005 at the operating frequency of 11.424 GHz. A snapshot of the electric field vector for a cross-section of the structure is shown on the right-hand-side of the figure.

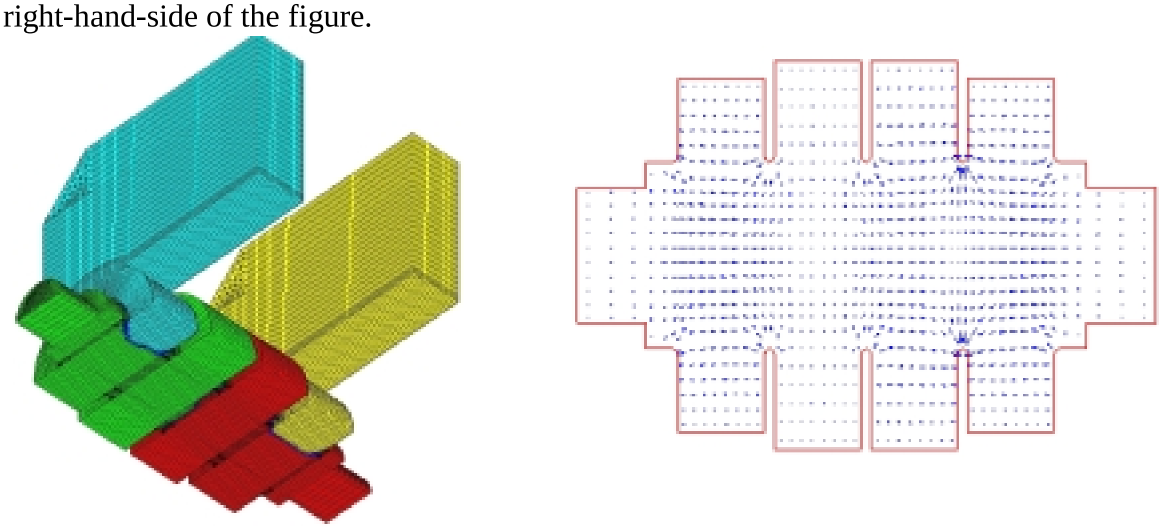

Figure 11: Simulation of an X-band input coupler showing a snapshot of the electric field vector, calculated using Tau3P, for a cross-section of the structure.

4 Plans for FY2000

Our goals for FY2000 involve a combination of supporting DOE Office of Science accelerator activities that require large-scale simulation, as well as laying the groundwork for our long-range goal of developing a reliable, maintainable, extensible terascale simulation capability for the accelerator community. In regard to the former, we will continue to support the NLC, SNS, and APT projects, as well as providing support to new projects requiring large scale simulation (for example, SPEAR-III).

Grand Challenge accomplishments such as the development of IMPACT, Omega3P, and Tau3P have already had a major impact, and will have a lasting presence, in the accelerator community. But the Grand Challenge team still has long-range goals that will require several years to complete. Our code-development activities for FY2000 are meant to have near-term applications to DOE projects as well helping us to reach our long-range goals. Planned activities for FY2000 are described below.

- **Beam Dynamics**
  Large scale beam dynamics simulations will reach their full potential only when it becomes possible to perform end-to-end simulations of complete accelerator systems. As a case-in-point, the above-mentioned SNS simulations of maximum particle amplitude are highly dependent on the initial state of the system. This is not surprising, since the majority of beam dynamics calculations are initial value problems. As such, we are now broadening our beam dynamics effort to include a wider range of accelerator systems. As mentioned previously, we have already begun to develop a parallel code for simulating beam dynamics in radio frequency quadrupoles (RFQs), which are the injectors found in most modern ion linacs. We will complete the development of this code in FY2000. Furthermore, while our emphasis up to now has been on linear accelerators, we now plan to continue development of the IMPACT code with the goal of extending it to treat circular accelerators as well. This involves adding a capability to treat bending magnets, for which we are well-positioned due to our use of split-operator methods which allow us to make use of existing bending magnet capability as found in Magnetic Optics. Using the same approach, we will also extend IMPACT to include the effects of nonlinear external focusing fields. Of more difficulty will be the treatment of space charge, which will involve a major effort because several models of space charge are needed to simulate a full accelerator system. As a case in point, consider a high-intensity proton system: At injection, typically in an RFQ, the space charge must be computed taking into account the nearby RFQ vanes that focus and accelerate the beam; in the linear accelerator sections that follow, the beam dimensions are of order 1 mm, which is small compared with the beam pipe radius, and each bunch can be treated as if it is isolated; if the beam is then injected into a storage ring, the beam becomes 10's of meters long, which is much larger than the pipe radius, and the space charge takes on an entirely different quality. We intend to build upon existing parallel software, and develop new software as needed, for the efficient, parallel computation of the space-charge fields in a variety of accelerator systems. We are already working with a European collaborator who is developing a new, tree-code-based space charge

solver that is built on the POOMA framework. This solver will be useful in a variety of situations, including modeling beam dynamics in cyclotrons.

Another activity aimed at broadening the applicability of our codes is to include a choice of particle-advance methods. While most beam dynamics codes use a coordinate as the independent variable, there are situations in which is necessary to use the time as the independent variable. As mentioned above, we have already begun to develop a parallel, 3D time-based code. Such a capability is needed, for example, to accurately model beam dynamics in high-brightness electron guns. The development of a 4^{th} generation light source will require very low emittance beams, and hence a predictive capability to design and optimize high-brightness guns using high-resolution, time-based particle simulations. Modeling of such systems will also require the inclusion of coherent synchrotron radiation effects.

In FY2000 we will continue our activities aimed at modeling noise and collisions in particle accelerators, and in developing methods to solve stochastic partial differential equations. One of our goals is to perform the first, fully self-consistent 3D Fokker-Planck simulations (i.e. simulations in which the damping and diffusion coefficients are treated from first principles). Such calculations are relevant to predicting the lifetime of beams in storage rings.
Finally, we intend to benchmark our IMPACT calculations by comparing simulation results with experimental results from a planned beam halo experiment to be carried out at the LANL Low Energy Demonstration Accelerator.

- **Electromagnetics**
  As mentioned above, the Tau3P and Omega3P codes developed so far under the Grand Challenge are already proving their usefulness to projects such as the NLC and SNS. However, there are important problems that can only be solved if new capabilities are added to the codes. In FY2000 we will begin to add two new capabilities in the form of (1) a complex-valued version of Omega3P and (2) a version of Tau3P that includes a rigid-beam model.

  The present version of Omega3P deals with loss-free cavities, so the eigenvalues of the associated sparse linear system are along the real axis. In most accelerating cavities the wall material, such as copper, has such low loss that it can be treated essentially as perfectly conducting. The wall loss power is then obtained as a perturbation to be calculated from the loss-free solution or the real eigenvector. Some cavities, however, are made of material with much higher loss. Others have lossy loads placed within them to damp higher order modes (HOM's). Many have apertures on the cavity wall which are connected to waveguides for external loading purposes. In all these cases where the damping effect can no longer be treated as a perturbation, the power loss, whether it be dissipative or diffractive in nature, has to be found self-consistently. This means the eigenvalue problem has to be formulated in the complex domain with the eigenvalues now residing in the complex plane. Following the finite element formulation of Omega3P, the matrix representation of Maxwell's equations with lossy materials reduces to the generalized complex symmetric eigenvalue problem. The latter is of great interest to the numerical analysis community because of its applicability in many areas of physics and engineering. However, unlike the real symmetric eigenvalue problem, the mathematical properties of the complex case are not well understood. As a result, many numerical algorithms have shown to converge poorly, especially when they are applied to large matrices. Preliminary work under the Grand Challenge revealed that an iterative scheme, using the loss-free solution as the starting vector, shows great promise in obtaining an accurate solution to the complex problem, even when the loss is significant. In FY2000 we plan to implement the algorithm into Omega3P and to benchmark the code against measured data from the SLAC B-factory (PEP-II) cavity and the NLC damped detuned structure. This will be a major undertaking because a new complex data structure is required and all the linear algebra routines will have to be rewritten for complex operations.

In FY2000 we will also begin to incorporate the effect of a rigid beam in Tau3P. The present version of the code was developed to simulate pulse propagation in transmission structures. Adding the effect of a rigid, transiting beam will enable the direct calculation of wakefield effects in accelerator components and systems. Wakefields are parasitic electromagnetic fields generated by a transiting beam through its interaction with the accelerator environment. If left uncontrolled, they can have adverse effects on the operation and performance of an accelerator. For example, short-range wakefields can lead to bunch lengthening that degrades the luminosity at the collision point. Long-range wakefields can cause unstable coupled bunch motion in a storage ring, and can also dilute the beam emittance of a bunch train travelling down a linac. Another effect of wakefields is the excessive heating of a beamline component due to the repeated deposition of HOM power by the beam through its excitation of trapped modes. These and other undesirable consequences can be prevented if a careful study of the wakefields in the beamline components is carried out. Wakefield analysis is already important in accelerator R&D but it takes on a particular urgency in future machines which plan to operate at higher currents and with smaller bunches. The development of a new version of Tau3P will involve considerable effort devoted to meshing, noise filtering and boundary condition implementation.

- **Combined Beam Dynamics and Electromagnetics**
  It is a long-range goal of the Grand Challenge team to develop a capability to model systems that involve strong coupling between beam dynamics and electromagnetics. An example is the output structure of a high-power klystron. Such a system is enormously complicated, and for this reason it can be argued that it has never been simulated correctly, and that no quantitative, first-principles predictive capability exists for such devices. (This explains the great difficulty that has been encountered worldwide in recent years in “pushing the envelope” in high-power klystron development.) First, the beam itself excites fields in the cavity (typically an “extended interaction” cavity), which act back on the beam. Second, at the same time, the fields are being extracted from the cavity itself. Third, all this occurs in an

extremely complicated 3D geometry, with a variety of boundary conditions, and possibly the use of absorbing materials to prevent the growth of high-order modes in the cavity. We intend to model such a system by treating it essentially as an electromagnetics problem with particles. (The alternative, to treat it as a beam dynamics problem with an electromagnetic model built into it, is difficult to do in a controlled and systematic way that converges sufficiently rapidly, and furthermore it is less accurate than the electromagnetics-oriented approach). This activity will require the use of hybrid meshes, i.e. those involving both regular and irregular components. Although valuable experience has been gained during the Grand Challenge with generating quality irregular meshes for pulse propagation, the inclusion of a beam poses new challenges to meshing in several respects. It is desirable, for example, that the mesh be uniform in the region of the beam in order to simplify the charge deposition and field interpolation related to particle propagation. (On a related note, a hybrid mesh that is uniform in the direction of beam travel will be needed in the new version of Tau3 in order to compute the wakefield in a straightforward way.) For the mesh to also conform to the irregular boundary walls, the mesh distribution has to consist of a uniform core surrounding the beam axis and an outer non-uniform region that connects the surface of the inner core to the outside walls. The situation is especially challenging in a klystron output structure, because some particles are decelerated to such an extent that they strike the walls of the output structure; hence they will move from the regular grid to the irregular grid, and will have to be treated accordingly. The ability to model systems involving strong coupling between beam dynamics and electromagnetics will require a capability to generate hybrid meshes of good quality, and to propagate particles through those hybrid meshes, but at the same time not imposing unrealistic compute resources for those computations.

5 Presentations at Technical Meetings

1. R. Ryne et al., "Prediction of Beam Halo Using Parallel Supercomputers," presented at the 7th Workshop on High-Intensity High-Brightness Hadron Beams, Lake Como, Wisconsin, September 1999.
2. K. Ko et al., "Large Scale Electromagnetic Modeling of Accelerator Structures and Components using High Performance Computers," Invited talk, 1999 Particle Accelerator Conference, New York, NY, March 1999.
3. R. Ryne et al., "The U.S. DOE Grand Challenge in Computational Accelerator Physics," Invited talk, XIX International Linac Conference, Chicago, IL, August 1998.
4. C.-K. Ng, B. McCandless, Y. Sun, M. Wolf, and K. Ko, "Tau3P: A Parallel Time-Domain Solver for the DOE Grand Challenge," 1999 International Particle Accelerator Conference, New York, NY, March 1999.
5. B. McCandless, Z. Li, Y. Sun, and K. Ko, "Omega3P: Modeling Next-Generation Particle Accelerators," 1999 Particle Accelerator Conference, New York, NY, March 1999.
6. J. DeFord and K. Ko, "3D Optimization Using a Client/Server Software Topology," 1999 Particle Accelerator Conference, New York, NY, March 1999.

7. I. Hofmann, J. Qiang, and R.D. Ryne, “Coherent Coupling Criterion for Three-Dimensional Halo Formation,” 1999 Particle Accelerator Conference, New York, NY, March 1999.
8. J. Qiang, R.D. Ryne, and S. Habib, “Parallel Object-Oriented Design in Fortran for Beam Dynamics Simulation,” 1999 Particle Accelerator Conference, New York, NY, March 1999.
9. J. Qiang, R.D. Ryne, and S. Habib, “Beam Halo Studies Using a Three-Dimensional Particle-Core Model,” 1999 Particle Accelerator Conference, New York, NY, March 1999.
10. R.M. Jones, N.M. Kroll, R.H. Miller, T. Higo, K. Ko, Z. Li, R.D. Ruth, V. Srinivas, and J.W. Wang, “The Dipole Wakefield for a Rounded Damped Detuned Linear Accelerator with Optimized Cell-to-Manifold Coupling,” XIX International Linac Conference, Chicago, IL, August 1998.
11. A.V. Fedotov, R.L. Gluckstern, S. Kurennoy, and R.D. Ryne, “Halo Formation in 3D Bunches with Different Phase Space Distributions,” XIX International Linac Conference, Chicago, IL, August 1998.
12. J.J. Barnard, S.M. Lund, and R.D. Ryne, “Self-Consistent 3D Simulations of Longitudinal Halo in RF Linacs,” XIX International Linac Conference, Chicago, IL, August 1998.
13. B. McCandless et. al, “Omega3P, Modeling the Next Generation of Accelerator Structures,” Supercomputing’98, Orlando, Florida, November 7-13, 1998.
14. M. Saparov et. al, “Multithreaded Sparse Matrix Library with Dynamic Load Balancing,” Supercomputing’98, Orlando, Florida, November 7-13, 1998.
15. Y. Sun et. al, “A Hybrid Scheme to Improve Jacobi-Davidson Method on Eigenvalue Problems,” International Workshop on Accurate Solution of Eigenvalue Problems, Penn State University, University Park, PA, July 20-23, 1998.

6 Selected Publications

1. J. Qiang, R.D. Ryne, S. Habib, and V. Decyk, "An Object-Oriented Parallel Particle-In-Cell Code for Beam Dynamics Simulation in Linear Accelerators," accepted for publication in the Proceedings of SC'99, Portland, OR, November 1999.
2. P.S. McCormick, J. Qiang, and R.D. Ryne, “Visualizing High-Resolution Accelerator Physics,” IEEE Computer Graphics and Applications, 19(5), pp. 11-13, September 1999.
3. W. Humphrey, R. Ryne, T. Cleland, J. Cummings, S. Habib, G. Mark, and J. Qiang, "Particle Beam Dynamics Simulations Using the POOMA Framework," in Lecture Notes in Computer Science #1505, Springer-Verlag, Dec. 1998.
4. A.V. Fedotov, R.L. Gluckstern, S. Kurennoy, and R. Ryne, "Halo Formation in 3D Bunches with Different Phase Space Distributions," Phys. Rev. ST Accel. Beams 2:014201, 1999.
5. R.L. Gluckstern, A. Fedotov, S. Kurennoy, and R. Ryne, "Halo Formation in Three Dimensional Bunches," Phys. Rev. E 58:04, October 1998.

6. J. Qiang, R.D. Ryne, and S. Habib, “Object-Oriented Parallel Particle-In-Cell Code for Beam Dynamics Simulation in Linear Accelerators,” submitted to J. Comp. Phys.
7. J. Qiang, R.D. Ryne, and S. Habib, “Beam Halo Studies Using a Three-Dimensional Particle-Core Model,” submitted to Phys. Rev. E.
8. C. Adolphsen et al., “Wakefield and Beam Centering Measurements of a Damped and Detuned X-Band Accelerator Structure,” Proceedings of the 1999 Particle Accelerator Conference, New York, NY, March 1999
9. G. Bowden et al., “A Compact RF Power Coupler for the NLC Linac,” Proceedings of the 1999 Particle Accelerator Conference, New York, NY, March 1999
10. I. Hofmann, J. Qiang, and R.D. Ryne, "Coherent Coupling Criterion for Three-Dimensional Halo Formation," Proceedings of the 1999 Particle Accelerator Conference, New York, NY, March 1999
11. Z. Li et al., “RDDS Cell Design and Optimization for the NLC Linac,” Proceedings of the 1999 Particle Accelerator Conference, New York, NY, March 1999
12. R. Ryne et al., "The U.S. DOE Grand Challenge in Computational Accelerator Physics," Proceedings of the XIX International Linac Conference, Chicago, IL, August 1998.
13. J. Qiang and S. Habib, "A Second-Order Stochastic Leap-Frog Algorithm for Multiplicative Noise," (in preparation).